

Variational optimization of effective atom centered potentials for molecular properties

O. Anatole von Lilienfeld, Ivano Tavernelli, and Ursula RothlisbergerDaniel Sebastiani

Citation: *The Journal of Chemical Physics* **122**, 014113 (2005); doi: 10.1063/1.1829051

View online: <http://dx.doi.org/10.1063/1.1829051>

View Table of Contents: <http://aip.scitation.org/toc/jcp/122/1>

Published by the [American Institute of Physics](#)

COMPLETELY

REDESIGNED!



**PHYSICS
TODAY**

Physics Today Buyer's Guide
Search with a purpose.

Variational optimization of effective atom centered potentials for molecular properties

O. Anatole von Lilienfeld, Ivano Tavernelli, and Ursula Rothlisberger^{a)}
*Ecole Polytechnique Fédérale de Lausanne, Institut des Sciences et Ingénierie Chimiques,
EPFL-BCH, CH-1015 Lausanne, Switzerland*

Daniel Sebastiani
Max-Planck-Institut für Polymerforschung, Ackermannweg 10, D-55128 Mainz, Germany

(Received 2 August 2004; accepted 14 October 2004; published online 14 December 2004)

In plane wave based electronic structure calculations the interaction of core and valence electrons is usually represented by atomic effective core potentials. They are constructed in such a way that the shape of the atomic valence orbitals outside a certain core radius is reproduced correctly with respect to the corresponding all-electron calculations. Here we present a method which, in conjunction with density functional perturbation theory, allows to optimize effective core potentials in order to reproduce ground-state molecular properties from arbitrarily accurate reference calculations within standard density functional calculations. We demonstrate the wide range of possible applications in theoretical chemistry of such optimized effective core potentials (OECPs) by means of two examples. We first use OECPs to tackle the link atom problem in quantum mechanics/molecular mechanics (QM/MM) schemes proposing a fully automatized procedure for the design of link OECPs, which are designed in such a way that they minimally perturb the electronic structure in the QM region. In the second application, we use OECPs in two sample molecules (water and acetic acid) such as to reproduce electronic densities and derived molecular properties of hybrid (B3LYP) quality within general gradient approximated (BLYP) density functional calculations. © 2005 American Institute of Physics. [DOI: 10.1063/1.1829051]

I. INTRODUCTION

Many *ab initio* electronic structure calculations, especially when dealing with large systems or heavy atoms, exploit the frozen-core approximation: only the chemically relevant valence electrons are treated explicitly while an effective core potential (ECP) or pseudopotential accounts for the effects of the core electrons.^{1–3} Within the field of condensed matter physics the use of pseudopotentials is especially crucial for plane wave based calculations. While atomic pseudopotentials were introduced in an empirical formulation in 1970,⁴ the development of *ab initio* pseudopotentials has been going on since 1959.^{5–9} ECPs for heavy atoms, molecules, and condensed phase calculations have been subject to improvement ever since.^{10,11} In 1996, analytic (Gaussian-based) norm conserving separable dual-space pseudopotentials were introduced by Goedecker *et al.* for some first, second, and third row elements¹² and extended to all atoms up to Radon including relativistic effects in 1998.¹³ The parameters of these ECPs are generated for each atom by iteratively minimizing an atomic penalty functional, which expresses the deviations of ECP-generated Kohn–Sham (KS) valence orbitals from their (relativistic) all-electron counterparts. Here, we propose a method which exploits an analogous procedure in such a way that not only atomic but also molecular properties are included as a reference. Specifically, we propose to consider complex molecu-

lar properties as target quantities in the penalty functional. In turn, the transferability requirements of the atomic pseudopotential are relaxed in favor of ECPs optimized for generalized molecular properties. In principle, our approach can be seen as an atom-based version of the optimized effective potential (OEP) method in density functional theory (DFT) (Refs. 14–18) where an orbital dependent external effective potential is optimized. However, while the OEP scheme exploits the variational principle, we suggest to tune the ECPs directly such as to reproduce some molecular reference property. So far, we have studied properties such as energies of interaction, forces on the nuclei, and electron densities. Reproducing forces and binding energies at the equilibrium geometry, we could show that extended and calibrated optimized ECPs (OECPs) can cure the lack of dispersion forces in commonly used exchange-correlation (xc) functionals.¹⁹ In this study, we illustrate the generality of OECPs by addressing electron densities as molecular properties. Zhao and co-workers^{20–22} have shown the validity of this approach for single atoms by computing successfully KS orbitals and energies. Here, we apply the same idea to molecules and aim to demonstrate its usefulness by tackling other kinds of problems within DFT based electronic structure theory: First, we show that our approach can be used for the design of special link atoms for binding quantum to classical fragments in quantum mechanics/molecular mechanics (QM/MM) applications. Second, by the means of OECPs we will target to reconstruct a given reference electronic density computed with a certain amount of nonlocal Hartree–Fock (HF) ex-

^{a)}Electronic mail: ursula.rothlisberger@epfl.ch
URL: <http://lcbepec21.epfl.ch>

change, using the B3LYP functional,^{23,24} within a standard local general gradient approximated (GGA) xc-potential calculation, using BLYP.^{24,25}

Hybrid QM/MM calculations have attracted a lot of attention in recent years because of their ability to accurately treat at the quantum level (QM part) molecules or molecular fragments (containing hundreds of atoms) embedded in a large environment (many thousands of atoms) which is described classically (MM part).^{26,27} However, when the classical surrounding is covalently bound to the quantum region, the problem of how to link the QM part to the MM region is encountered. A popular choice is to cap all boundary atoms by hydrogens, however, this often leads to a distortion of the electronic density.²⁸ Several research groups proposed to design special capping effective potentials to reduce this effect.^{29–32} Recently, a general scheme to fit pseudopotential parameters in order to reproduce model system properties such as bond lengths or charges has also been introduced.³³ Alternatively, the frontier orbitals of the link atom which point into the direction of the MM region can be considered to be frozen.^{34,35} In the present study, we propose to optimize systematically the capping ECPs at the boundaries using density functional perturbation theory. Specifically, using perturbed electron densities we will show how to iteratively minimize differences in electron density between the QM region and an all-atom QM reference calculation. In this context, it should be mentioned that there are also frozen density approaches avoiding the linking atom problem by construction.³⁶

In the second application we attempt to improve the performance of GGA based DFT functionals using OECPs. Often, hybrid functionals such as B3LYP are able to yield a more accurate description of molecular systems³⁷ and reaction profiles³⁸ than calculations using the GGA such as, e.g., BLYP. Especially transition states and regions of low electron density, e.g., hydrogen bonds, are described significantly better when HF exchange is included in the xc functional.³⁹ Unfortunately, the inclusion of HF exact exchange, such as it is necessary for B3LYP DFT calculations, into the wave function optimization algorithm increases the scaling of the computational cost for plane wave calculations typically by one order of magnitude. As a solution to this problem we propose to use specially designed OECPs within the BLYP DFT level of theory in order to rebuild the electronic density from a corresponding B3LYP reference calculations of the same molecular system. Therefore, within this application, we will investigate how the OECP method responds to the following question: within a local BLYP calculation, how do those terms of the total external potential which are constituted by the nonlocal ECPs have to be modified in order to yield an electron density which resembles as close as possible the electron density corresponding to unmodified ECPs plus the nonlocal B3LYP xc potential? As molecular test systems for this application, we have chosen water and acetic acid, which are two ubiquitous solvent molecules and which can act simultaneously as hydrogen bond acceptors and donors.

II. METHODS AND COMPUTATIONAL DETAILS

The basic idea of the approach has been introduced in Ref. 19 and is outlined here in more detail. The method employs the iterative minimization of a penalty functional \mathcal{P} by variational tuning of a certain number of ECP parameters. The way the penalty functional is designed determines which molecular property will be optimized. Usually it represents the difference of selected physical properties, e.g., the electronic density and its higher moments, computed with a given set of values for the ECP parameters (which we denote by σ_i), with respect to a reference system which can be computed at any level of theory. Thus, the minimization of $\mathcal{P}(\{\sigma_i\})$ (which may depend on all quantities that can be expressed via the KS orbitals of the system) allows to approach as much as possible an arbitrary reference within the limits of the chosen functional form of the ECP approach. In principle, this approach can be applied to any molecular ground state property such as the electronic density $n(\mathbf{r})$, its multipole moments, ionic forces, as well as the total energy. For the applications of this study, we choose to let $\mathcal{P}(\{\sigma_i\})$ depend only indirectly on the set of parameters $\{\sigma_i\}$ by means of the electronic density.

Without any loss of generality, we illustrate the procedure supposing a simple functional form only dependent on the electronic density $n(\mathbf{r}) = \sum_{k=1}^N |\phi_k(\mathbf{r})|^2$, where N is the number of occupied KS orbitals ϕ_k :

$$\mathcal{P}[n(\mathbf{r}, \{\sigma_i\})] = \int d^3r w(\mathbf{r}) \mathcal{F}(n(\mathbf{r})). \quad (1)$$

The weighting function $w(\mathbf{r})$ can be used to restrict the integration to a certain volume or to certain atoms only. The minimization of Eq. (1) is performed by following the gradient of \mathcal{P} with respect to the ECP parameters σ_j ,

$$\frac{d\mathcal{P}}{d\sigma_j} = \int d^3r w(\mathbf{r}) \frac{\partial \mathcal{F}}{\partial n}(\mathbf{r}) \frac{dn(\mathbf{r})}{d\sigma_j}. \quad (2)$$

The derivatives $dn(\mathbf{r})/d\sigma_j$, which we will denote $n_j^{(1)}(\mathbf{r})$, represent the linear change in the electronic structure of the system which is induced by a variation of one of the ECP parameters $\sigma_j \mapsto \sigma_j + d\sigma_j$: $n_j^{(1)}(\mathbf{r}) = \sum_{i=1}^N [\phi_i^{*(0)}(\mathbf{r}) \phi_{ij}^{(1)}(\mathbf{r}) + \phi_{ij}^{*(1)}(\mathbf{r}) \phi_i^{(0)}(\mathbf{r})]$. This linear density response can be computed through density functional perturbation theory, where the perturbation Hamiltonian is given by the change in the ECP due to the variation of σ_j ,

$$\hat{\mathcal{H}}'_j = \frac{\partial \hat{V}^{\text{ECP}}(\{\sigma_i\})}{\partial \sigma_j}. \quad (3)$$

We have implemented Eq. (3) in the program code CPMD.⁴⁰ We apply Eq. (3) for each orbital within the implementation of a general variational density functional perturbation theory module⁴¹ to determine the vector of response densities $\{n_j^{(1)}(\mathbf{r})\}$. Following the gradients of Eq. (2) until convergence of the penalty \mathcal{P} [Eq. (1)] we determine the set of ECP parameters $\{\sigma_i\}$, which represent best the OECP corresponding to the chosen penalty. To implement Eq. (3) we use analytical ECPs as the one introduced by Goedecker *et al.*¹²

They consist of a local $V^{\text{loc}}(\mathbf{r})$ and of a sum of nonlocal parts $V_l^{\text{nl}}(\{\sigma_j\}, \mathbf{r}, \mathbf{r}')$ with specific projectors for each angular momentum channel l ,

$$V_1^{\text{ECP}}(\mathbf{r}, \mathbf{r}') = V^{\text{loc}}(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') + \sum_l V_l^{\text{nl}}(\mathbf{r}, \mathbf{r}'), \quad (4)$$

$$V^{\text{loc}}(\mathbf{r}) = \frac{-Z_{\text{ion}}}{r} \operatorname{erf}\left[\frac{r}{r_{\text{loc}}\sqrt{2}}\right] + \exp\left[-\frac{r^2}{2r_{\text{loc}}^2}\right] \times \left[C_1 + C_2 \left(\frac{r}{r_{\text{loc}}}\right)^2 + C_3 \left(\frac{r}{r_{\text{loc}}}\right)^4 + C_4 \left(\frac{r}{r_{\text{loc}}}\right)^6 \right], \quad (5)$$

$$V_l^{\text{nl}}(\mathbf{r}, \mathbf{r}') = \sum_{m=-l}^{+l} Y_{lm}(\hat{\mathbf{r}}) \sum_{j,h=1}^3 p_{lh}(r) h_{lhj} p_{lj}(r') Y_{lm}^*(\hat{\mathbf{r}}'), \quad (6)$$

where $p_{lh}(r) \propto r^{l+2(h-1)} \exp(-r^2/(2r_l^2))$, $r = |\mathbf{r} - \mathbf{R}_l|$ on the position \mathbf{R}_l of nucleus l , $\hat{\mathbf{r}}$ is the unit vector in the direction of \mathbf{r} , and Y_{lm} denotes a spherical harmonic. The parameters $\{r_{\text{loc}}, C_1, C_2, C_3, C_4, h_{lhj}, r_l, \dots\}$ span the parameter space $\{\sigma_j\}$, its dimensionality being determined by the largest angular momentum component of the ECP. We could also have investigated the effect of optimizing additional larger angular momentum channels, which would not be necessary in conventional ECPs because for carbon or oxygen the ECP series can be safely truncated at the first l term not contained in the core. However, in this study we have chosen to restrict the optimization of the ECPs in all cases to the already existing parameters defining an ECP of a second row atom such as carbon or oxygen. In our dispersion forces related study¹⁹ we have used for optimization only parameters defining an additional angular momentum channel, which was introduced solely for the purpose of the optimization. We present the explicit form of Eq. (3) in the Appendix.

For the two applications described in detail in the following section, we have chosen to optimize electron densities as molecular properties. In full analogy to Eq. (22) in Ref. 21 we have chosen the following penalty functional:

$$\mathcal{P}[n(\mathbf{r})] = \int d^3r w(\mathbf{r}) |n^{\text{ref}}(\mathbf{r}) - n(\mathbf{r})|^2 \quad (7)$$

with individual weighting functions $w(\mathbf{r})$ and appropriate reference densities $n^{\text{ref}}(\mathbf{r})$. According to Eq. (2) the gradient of Eq. (7) with respect to an ECP parameter σ_j is given by

$$\frac{d\mathcal{P}[n(\mathbf{r})]}{d\sigma_j} = 2 \int d^3r w(\mathbf{r}) n_j^{(1)}(\mathbf{r}) [n(\mathbf{r}) - n^{\text{ref}}(\mathbf{r})]. \quad (8)$$

All calculations were carried out using CPMD (Ref. 40) at an orbital plane wave cutoff of 100 Ry and using the isolated system Poisson solver according to Ref. 42.

III. RESULTS AND DISCUSSION

A. QM/MM link atom optimization

As a first application, we address the problem of designing optimal intramolecular link atoms between the quantum and classical regions in QM/MM simulations where typically a C–C bond is cut. The substitution of a real carbon atom

TABLE I. Calculated total permanent dipole moments $|\mu|$ for R–CO₂H·R = CH₃ corresponds to full QM acetic acid, D^{opt} is a link OECP replacing the CH₃ group, H indicates hydrogen capping, D^{con} is an empirically optimized monovalent link carbon atom which reproduces a H₃C–CH₃ bond length (Ref. 32), and F corresponds to fluorine capping. All values are in debye.

Pro\R	CH ₃	D ^{opt}	H	D ^{con}	F
$ \mu $	1.70	1.32	0.69	3.10	2.31

within a macromolecule by a so-called link atom, designed to saturate the open valence, generally leads to a distortion of the electronic density in the QM neighborhood of the link atom. We want to investigate if we can minimize this perturbing effect of the artificial QM/MM bond cleavage on the electronic density in the QM region by applying our procedure to identify an OECP for the link atom.

As a model system for the link atom optimization we have chosen acetic acid where the methyl group is replaced by a link OECP, and the weighting function $w(\mathbf{r})$ is chosen such as to restrict the integration volume in Eq. (7) to the union of spheres of a radius of 1.5 Å (corresponding roughly to a C–C bond length) centered on each atom, while on the linking nucleus the corresponding radius is only 0.7 Å. By doing so, we exclude from the penalty functional $\mathcal{P}^{\text{link}}$ the necessarily different electron density of the methyl group while keeping the rest of the QM region in the integration volume.

Using a gradient based optimization technique, we have minimized $\mathcal{P}^{\text{link}}$ in the space of all ECP parameters of the capping atom to identify the OECP. The results for calculated total permanent dipole moments are presented in Table I. We have investigated acetic acid using four different replacements for the methyl group: an optimized heptavalent ECP, hydrogen capping, an empirically modified monovalent link atom³² based on a carbon ECP, and fluorine capping. Subsequently, we have optimized the molecular geometry using the four alternative saturation approaches. The C–D^{opt} bond length is increased negligibly in the case of the OECP, while hydrogen or fluorine capping result in significant changes in the bond lengths. In comparison to the full QM calculation of acetic acid, the OECP D^{opt} closely reproduces the total dipole moment, while all alternatives yield rather unsatisfying results. The small dipole moment in the case of hydrogen saturation reflects the fact that the hydrogen 1s electron is less polarizable by the electron attracting–CO₂H group than the methyl group. In contrast to that, when using a conventional empirically adjusted monovalent linking carbon atom or fluorine capping the dipole moment is overestimated and inverted. Especially a wrong dipole moment often constitutes one of the major sources of errors at the boundary of QM/MM calculations. Furthermore, an accurate description of the QM/MM boundary makes it possible to reduce the size of the QM region, thereby allowing a decrease in computational cost. The scheme presented here can easily be implemented into an automatic procedure for the generation of optimal link ECPs for arbitrary chosen reference situations.

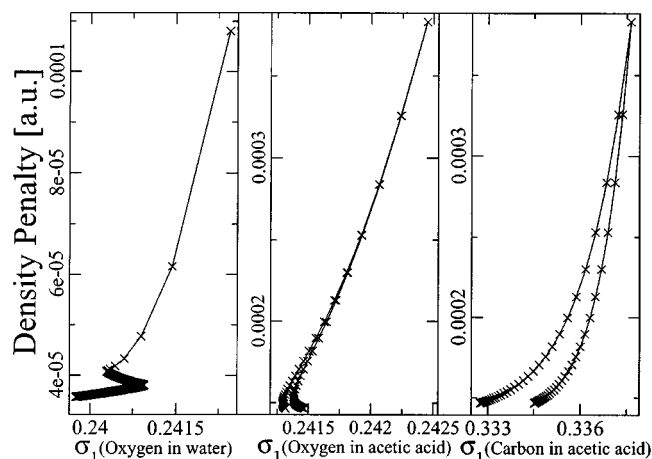


FIG. 1. History of the minimization of the density penalty as a function of σ_1 of oxygen in water and acetic acid and of carbon in acetic acid.

B. B3LYP density reconstruction using BLYP

Using our approach, we have also determined OECP parameters for all heavy atoms in water and acetic acid which yield best the hybrid B3LYP electron density within a pure DFT BLYP functional calculation. They are able to reproduce the electronic density resulting from the B3LYP reference calculations with the smallest least square error. In principle, we could have taken as a reference an electron density originating from any level of theory, such as, e.g., Møller-Plesset or coupled cluster calculations. Instead of a reference electron density one could also use other molecular properties, such as, e.g., ionic forces (geometry) or KS eigenvalues for the construction of the penalty functional. Here, we use the simplest form of a penalty functional, Eq. (7), which imposes an approximate reconstruction of the electron density obtained from a B3LYP calculation, namely, $n^{ref} = n_{B3LYP}$. We have taken the ECP parameters of all oxygen and carbon atoms as variables and minimized Eq. (7) in the original B3LYP-optimized geometry.

For the oxygen and carbon ECPs in water and acetic acid, the progressive minimization of the penalty in its high-dimensional parameter space is shown as a function of only σ_1 [$= r_{loc}$, Eq. (5)] in Fig. 1. σ_1 represents approximately the core radius of the ECP. It is interesting to note that there seems to be a general trend to decrease σ_1 in order to reproduce the B3LYP xc functional better, independently of the atom type or the molecule. Throughout the minimization, while also all the other ECP parameters (not shown in Fig. 1) vary only by a few percents, \mathcal{P} decreases to roughly one third of its initial value.

Furthermore, we would like to point out that—as one would expect—the optimized parameters of the same atom type in different chemical environments adopt in a slightly different manner, relaxing thereby the criterion of transferability. For example, the OECP of a carbon atom in a carboxylic group or in a methyl group differ. However, as presented in Fig. 1, the final core radii σ_1 of the two carbon or oxygen atoms in acetic acid differ only by less than 1%.

In Tables II and III, we present for comparison the deviations of several quantities which are derived from the

TABLE II. Deviation Δ with respect to the B3LYP results for calculated total dipole moments $|\mu|$, RESP charges q^{RESP} , bond lengths d_{O-H} , and angles for water. Dipole moments are in debye, charges are in atomic units, and angles are in degrees.

Method	$\Delta \mu $	Δq_O^{RESP}	Δq_H^{RESP}	Δd_{O-H}	$\Delta \angle_{HOH}$
BLYP+ECP ^a	0.038	0.022	0.011	0.018	0.9
BLYP+OECP	0.010	0.009	0.004	0.016	0.3

^aECPs taken from Goedecker *et al.* (Ref. 12).

electronic density with respect to the corresponding B3LYP results. We have computed dipole moments, atomic restrained electrostatic potential derived charges [RESP (Ref. 43)] and geometries of water and acetic acid.

We observe a systematic tendency for all considered properties to approach the B3LYP values. With respect to the B3LYP results, electronic properties such as dipole moments and RESP charges improve in general by an order of magnitude. Interestingly, the RESP charge on the carboxylic carbon atom does not change. We suspect that this exclusive behavior is due to the somewhat more complex electronic structure surrounding this carbon atom having three σ and one π bond. Also geometrical distances are closer to the B3LYP values but they do not improve to the same extent as the electronic properties. But to account fully for properties other than those which are directly derived from the electron density, a more suitable penalty needs to be defined which includes, e.g., the ionic forces explicitly.

Here, we would like to point out that due to the significant improvement in the description of the charge distribution obtained using OECPs together with a GGA xc functional, also a more accurate description of hydrogen bonds should be possible. We have therefore computed the interaction energy ($E^{int} = E^{Dimer} - 2E^{Monomer}$) of the water and of the acetic acid dimers. The deviation with respect to the B3LYP results decreases in the case of the water dimer from 0.38 to 0.20 kcal/mol and in the case of the acetic acid dimer even from 0.97 to 0.03 kcal/mol.

The computational advantage of the use of the OECP-BLYP scheme instead of the ECP-B3LYP scheme lies in the computational cost, which within plane wave calculations is decreased by one order of magnitude when BLYP is used instead of B3LYP.⁴⁴

IV. CONCLUSIONS

We have presented an iterative method to tune on a molecular level parameters of analytic ECPs in electronic structure calculations by means of minimizing a suitably defined

TABLE III. Deviation Δ with respect to the B3LYP results for calculated total dipole moments $|\mu|$, RESP charges q^{RESP} , and the O–H bond length d_{O-H} for acetic acid: $H_3C^{(1)}C^{(2)}O^{(1)}O^{(2)}H^{(1)}$. Dipole moments are in debye, and charges and distances are in atomic units.

Method	$\Delta\mu_z$	$\Delta q_{C^{(1)}}^{RESP}$	$\Delta q_{C^{(2)}}^{RESP}$	$\Delta q_{O^{(1)}}^{RESP}$	$\Delta q_{O^{(2)}}^{RESP}$	$\Delta q_{H^{(1)}}^{RESP}$	Δd_{O-H}
BLYP+ECP ^a	0.016	0.011	0.016	0.011	0.016	0.011	0.017
BLYP+OECP	0.002	0.001	0.017	0.006	0.005	0.000	0.012

^aECPs taken from Goedecker *et al.* (Ref. 12).

penalty functional. The penalty consists of an appropriate and differentiable function of the electronic density, and is constructed in such a way as to penalize deviations from a given reference calculation. This approach has been implemented in the plane wave pseudopotential code CPMD.⁴⁰

We have illustrated the potential of the method by means of two applications: we have constructed an OECP for a link atom for bond cuts in QM/MM schemes, which minimizes the perturbation of the QM region using an electron density penalty to quantify the differences between a complete QM calculation and a QM/MM calculation. The error on the quantum region introduced by the use of link atoms can thus be generally minimized within this scheme. The minimization of the error has been demonstrated by obtaining a correct dipole moment even when the QM/MM boundary is included. The approach is general enough to be used in order to obtain systematically the optimal capping OECP for any kind of bond by minimizing the above mentioned penalty inside the quantum region. Therefore, using our OECPs, the polarity of the QM/MM boundary can be conserved, which is important for the accuracy of the QM/MM approach. Furthermore, it is conceivable to reduce the size of the QM region because of the increased accuracy of the description of the QM/MM bond, thus allowing a decrease in computational cost. Hence, the scheme presented here is an automatic procedure to generate optimal link ECPS.

As a second application an electronic density computed using the B3LYP xc functional has been reconstructed within an OECP-BLYP calculation. We could show significant improvements for electrostatic properties such as dipole moments, RESP charges, and dimer interactions. This suggests that any kind of electron density corresponding to a given potential of arbitrary complexity can be approximately reconstructed by using our scheme of tuning the external atom centered nonlocal potential represented by the OECPs. This is of interest for the well known quest for better approximations to the unknown yet existing exact local exchange-correlation potential. We are pursuing further calculations of this kind with more accurate reference methods, such as explicitly correlated approaches. Considering the decrease of computational cost, in the special case of B3LYP the use of OECPs seems to be very promising to carry out more accurate “B3LYP-emulating” condensed phase plane wave calculations. However, the assessment of the transferability of these OECPs with respect to other atoms or dynamical influences is beyond the scope of this study and will be published elsewhere.

We conclude that with the use of OECPs in molecules the external potential can be adopted to special purposes. The inherent technical advantage of our approach consists in the fact that the computational cost is not increased because the standard pseudopotential machinery as it is implemented in most plane wave electronic structure codes can be used. Since molecules consist of atoms it is also legitimate to use series of atom based optimized potentials as a correction to the external potential to meet additional requirements on the description of the molecule.

This general scheme can also be applied to other purposes by using other references, more complex tailor-made

penalty functionals, and extended functional forms of the OECPs.

ACKNOWLEDGMENTS

The authors are grateful for discussions with A. Laio, P. Cummins, M. Sprik, T. Wesolowski, G. von Freymann, and M. Parrinello. A.v.L. acknowledges support from the Swiss National Science Foundation, Grant No. 510573.

APPENDIX: PERTURBATION HAMILTONIAN

Here, we give the explicit form of the perturbation Hamiltonian in Eq. (3) for a given Goedecker pseudopotential¹² V^{ECP} of maximal angular momentum $l = s$. \hat{H}'_j is the analytic derivative of the real space form of the analytic nonlocal and separable pseudopotential in Eqs. (5)–(6) with respect to the j th parameter σ_j . The parameters $\{\sigma_j\}$ are defined in the following order: $\{r_{\text{loc}}, C_1, C_2, r_s, h_{s11}\}$. $\hat{H}'|\phi(\mathbf{r})\rangle = \sum_j \hat{H}'_j|\phi(\mathbf{r})\rangle = \sum_j \partial_{\sigma_j} V^{\text{ECP}}|\phi(\mathbf{r})\rangle$, where the individual components are given by

$$\frac{\partial}{\partial \sigma_1} V^{(\text{loc})}(\mathbf{r}) \phi(\mathbf{r}) = \left[\frac{Z_{\text{ion}} \sqrt{2}}{\sqrt{\pi} \sigma_1^2} + \frac{r^2 \sigma_3}{\sigma_1^3} \left(\frac{r^2}{\sigma_1^2} - 2 \right) \right] \times \exp \left[-\frac{r^2}{2\sigma_1^2} \right] \phi(\mathbf{r}),$$

$$\frac{\partial}{\partial \sigma_2} V^{(\text{loc})}(\mathbf{r}) \phi(\mathbf{r}) = \exp \left[-\frac{r^2}{2\sigma_1^2} \right] \phi(\mathbf{r}),$$

$$\frac{\partial}{\partial \sigma_3} V^{(\text{loc})}(\mathbf{r}) \phi(\mathbf{r}) = \frac{r^2}{\sigma_1^2} \exp \left[-\frac{r^2}{2\sigma_1^2} \right] \phi(\mathbf{r}),$$

$$\begin{aligned} \frac{\partial}{\partial \sigma_4} V_i^{(\text{nl})}(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}) &= \frac{\partial}{\partial \sigma_4} \int d\mathbf{r}' V_i^{(\text{nl})}(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}') \\ &= \left[\frac{1}{\pi^{3/2} \sigma_5^3} \int d\mathbf{r}' \exp \left[-\frac{r'^2}{2\sigma_5^2} \right] \phi(\mathbf{r}') \right] \exp \left[-\frac{r^2}{2\sigma_5^2} \right], \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial \sigma_5} V_i^{(\text{nl})}(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}) &= \left[\left(\frac{\sigma_4 r^2}{\sigma_5^6 \pi^{3/2}} - \frac{3\sigma_4}{\sigma_5^4 \pi^{3/2}} \right) \int d\mathbf{r}' \exp \left[-\frac{r'^2}{2\sigma_5^2} \right] \phi(\mathbf{r}') \right. \\ &\quad \left. + \frac{\sigma_4}{\sigma_5^6 \pi^{3/2}} \int d\mathbf{r}' r' \exp \left[-\frac{r'^2}{2\sigma_5^2} \right] \phi(\mathbf{r}') \right] \exp \left[-\frac{r^2}{2\sigma_5^2} \right]. \end{aligned}$$

$r = |\mathbf{r} - \mathbf{R}_I|$ on the position \mathbf{R}_I of nucleus I. $\phi(\mathbf{r})$ is a KS orbital.

¹H. Hellmann, J. Chem. Phys. **3**, 61 (1935).

²H. Hellmann, J. Chem. Phys. **4**, 324 (1936).

³U. von Barth and C. D. Gelatt, Phys. Rev. B **21**, 2222 (1980).

⁴V. Heine and D. Weaire, Solid State Phys. **24**, 249 (1970).

⁵G. B. Bachelet, D. R. Hamann, and M. Schluter, Phys. Rev. B **26**, 4199 (1982).

- ⁶J. C. Phillips and L. Kleinman, *Phys. Rev.* **116**, 287 (1959).
- ⁷J. D. Weeks and S. A. Rice, *J. Chem. Phys.* **49**, 2741 (1968).
- ⁸G. B. Bachelet, D. R. Hamann, and M. Schluter, *Phys. Rev. Lett.* **43**, 1494 (1979).
- ⁹P. A. Christiansen, Y. S. Lee, and K. S. Pitzer, *J. Chem. Phys.* **71**, 4445 (1979).
- ¹⁰W. E. Pickett, *Comput. Phys. Rep.* **9**, 115 (1989).
- ¹¹M. Dolg, in *Modern Methods and Algorithms of Quantum Chemistry*, edited by J. Grotenдорst (John von Neumann Institute for Computing, Jülich, NIC Series, 2000), Vol. 3, p. 507.
- ¹²S. Goedecker, M. Teter, and J. Hutter, *Phys. Rev. B* **54**, 1703 (1996).
- ¹³C. Hartwigsen, S. Goedecker, and J. Hutter, *Phys. Rev. B* **58**, 3641 (1998).
- ¹⁴R. T. Sharp and G. K. Horton, *Phys. Rev.* **90**, 317 (1953).
- ¹⁵A. Görling and M. Levy, *Phys. Rev. A* **50**, 196 (1994).
- ¹⁶L. Fritsche and J. M. Yuan, *Phys. Rev. A* **57**, 3425 (1998).
- ¹⁷R. Colle and R. K. Nesbet, *J. Phys. B* **34**, 2475 (2001).
- ¹⁸W. Yang and Q. Wu, *Phys. Rev. Lett.* **89**, 143002 (2002).
- ¹⁹O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani, *Phys. Rev. Lett.* **93**, 153004 (2004).
- ²⁰Q. Zhao and R. G. Parr, *Phys. Rev. A* **46**, 2337 (1992).
- ²¹Q. Zhao and R. G. Parr, *J. Chem. Phys.* **98**, 543 (1993).
- ²²Q. Zhao, R. C. Morrison, and R. G. Parr, *Phys. Rev. A* **50**, 2138 (1994).
- ²³A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- ²⁴C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- ²⁵A. D. Becke, *J. Chem. Phys.* **88**, 2547 (1988).
- ²⁶A. Warshel, in *Computer Modeling of Reactions in Enzymes* (Wiley, New York, 1991).
- ²⁷ACS, in *Combined Quantum Mechanical and Molecular Mechanical Methods*, edited by J. Gao and M. A. Thompson (American Chemical Society, Washington, DC, 1999).
- ²⁸N. Reuter, A. Dejaegere, B. Maigret, and M. Karplus, *J. Phys. Chem. A* **104**, 1720 (2000).
- ²⁹D. Bakowies and W. Thiel, *J. Phys. Chem.* **100**, 10580 (1996).
- ³⁰Y. Zhang, T.-S. Lee, and W. Yang, *J. Phys. Chem.* **110**, 46 (1999).
- ³¹F. Bessac, F. Alary, Y. Carissan, J. L. Heully, J. P. Daudey, and R. Poteau, *J. Mol. Struct.: THEOCHEM* **632**, 43 (2003).
- ³²M. Sulpizi, U. Rothlisberger, A. Laio, A. Cattaneo, and P. Carloni, *Bio-phys. J.* **82**, 359 (2002).
- ³³G. A. DiLabio, M. M. Hurley, and P. A. Christiansen, *J. Chem. Phys.* **116**, 9578 (2002).
- ³⁴A. Warshel and M. Levitt, *J. Mol. Biol.* **103**, 227 (1976).
- ³⁵G. Monard, M. Loos, V. Théry, K. Baka, and J.-L. Rivail, *Int. J. Quantum Chem.* **58**, 153 (1996).
- ³⁶T. A. Wesolowski and A. Warshel, *J. Phys. Chem.* **97**, 8050 (1993).
- ³⁷L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, *J. Chem. Phys.* **109**, 42 (1998).
- ³⁸B. G. Johnson, C. A. Gonzales, P. M. W. Gill, and J. A. Pople, *Chem. Phys. Lett.* **221**, 100 (1994).
- ³⁹W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory* (Wiley-VCH, 2002).
- ⁴⁰J. Hutter *et al.*, Computer code CPMD, version 3.8 (IBM Corp., Stuttgart, 1990), <http://www.cpmid.org>
- ⁴¹A. Putrino, D. Sebastiani, and M. Parrinello, *J. Chem. Phys.* **113**, 7102 (2000).
- ⁴²G. Martyna and M. Tuckerman, *J. Chem. Phys.* **110**, 2810 (1999).
- ⁴³C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman, *J. Phys. Chem.* **97**, 10269 (1993).
- ⁴⁴Using CPMD (Ref. 40) on one CPU of an AMD OPTERON (244 1.8 GHz 4 GB memory) work station an average wave function optimization step of acetic acid in a $8 \times 8 \times 9 \text{ \AA}^3$ box at a wave function cutoff of 100 Ry took ≈ 80 s using B3LYP and ≈ 7.8 s using BLYP.